

# The Greedy Matching Algorithm in Random Regular Graphs and Hypergraphs

Deepak Bal\*

Patrick Bennett†

## Abstract

In this note, we analyze the performance of the greedy matching algorithm in sparse random graphs and hypergraphs with fixed degree sequence. We use the differential equations method and apply a general theorem of Wormald. The main contribution of the paper is an exact solution of the system of differential equations. In the case of  $k$ -uniform,  $\Delta$ -regular hypergraphs this solution shows that the greedy algorithm leaves behind

$$\left( \frac{1}{(k-1)(\Delta-1)} \right)^{\frac{\Delta}{(k-1)(\Delta-1)-1}} + o(1)$$

fraction of the vertices.

## 1 Introduction

A **matching** in a hypergraph is a collection of vertex disjoint edges. All hypergraphs considered in this note are  $k$ -uniform with some  $k \geq 2$ . The algorithmic theory of matchings in ordinary graphs ( $k = 2$ ) is very well studied. In particular, Edmond's Blossom Algorithm provides a polynomial time algorithm to find the largest size matching in a general graph (see e.g. [10]). However, the problem of finding the largest matching in a  $k$ -uniform hypergraph is NP-complete for  $k \geq 3$ .

One of the most basic algorithms for finding a maximal (not necessarily maximum) matching in a hypergraph is the random greedy algorithm. In each step of the algorithm, a uniformly random edge is selected to be added to the matching and all edges which intersect this edge are deleted from the host graph.

The performance of this algorithm was analyzed on arbitrary (ordinary) graphs by Dyer and Frieze [6], on dense random graphs by Tinhofer [13], and on sparse random graphs by Dyer, Frieze and Pittel [7]. The analysis in [7] was extended to sparse random hypergraphs in the Ph.D. thesis of Chebolu [4].

The random (hyper)graphs considered in [7] and [4] are drawn uniformly among all graphs with  $n$  vertices and a  $m = cn$  many edges. In this paper, we also consider sparse random graphs, but those with a fixed degree sequence.

---

\*Department of Mathematical Sciences, Montclair State University, Montclair, NJ, 07043 U.S.A.  
 deepak.bal@montclair.edu

†Department of Mathematics, Western Michigan University, Kalamazoo, MI, 49008 U.S.A.  
 patrick.bennett@wmich.edu

**Input** : Hypergraph  $H = (V, E)$   
**Output**: Matching  $M$   
 $M = \emptyset$ ;  
**while**  $E \neq \emptyset$  **do**  
    Select  $e \in E$  uniformly at random;  
     $M \leftarrow M \cup \{e\}$ ;  
     $E \leftarrow E \setminus \{e' \in E : e' \cap e \neq \emptyset\}$ ;  
**end**  
**return**  $M$ ;

**Algorithm 1:** RANDOM GREEDY

Suppose  $\mathbf{z} = (\zeta_1, \zeta_2, \dots, \zeta_\Delta) \in \mathbb{R}_+^\Delta$  is a fixed vector with  $\zeta_\Delta > 0$  and  $\sum_{i=1}^\Delta \zeta_i = 1$  ( $\mathbb{R}_+$  represents the non-negative real numbers). Let  $\mathcal{H}(n, k, \mathbf{z})$  represent the probability space with uniform distribution over all  $k$ -uniform hypergraphs on  $n$  vertices with  $\zeta_i n$  many vertices of degree  $i$  (we assume the degree sum over vertices is divisible by  $k$  and omit floors and ceilings). In the special case when  $\zeta_\Delta = 1$ ,  $\mathcal{H}(n, k, \mathbf{z})$  represents the random  $k$ -uniform,  $\Delta$ -regular hypergraph and we denote this by  $\mathcal{H}(n, k, \Delta)$ . Our main result is the following.

**Theorem 1.1.** *Suppose  $k, \Delta \geq 2$ , and define the functions*

$$P(x) := \sum_{i=1}^\Delta \zeta_i x^i, \quad Q(x) := \int_x^1 \frac{(k-1)wP''(w)}{P'(w)^k} dw,$$

*and let  $0 < \alpha_{\text{end}} < 1$  be the unique root of the equation*

$$\alpha = Q(\alpha)P'(\alpha)^{k-1}.$$

*Algorithm RANDOM GREEDY run on  $H \sim \mathcal{H}(n, k, \mathbf{z})$  produces a matching which covers all but*

$$[P(\alpha_{\text{end}}) + o(1)] n$$

*many vertices whp*<sup>1</sup>.

Note that since  $P(x)$  is just a polynomial, the antiderivative in the definition of  $Q(x)$  can (at least in principle) be calculated using partial fractions. Unfortunately, in general this antiderivative will be messy and involve logarithms and arctangents, in which case one would probably resort to numerical methods to approximate the root of the equation  $\alpha = Q(\alpha)P'(\alpha)^{k-1}$ . However, the solution can be written explicitly in the case corresponding to regular hypergraphs:

**Corollary 1.2.** *Suppose that  $k \geq 2$  and  $\Delta \geq 2$  are integers such that  $k + \Delta \geq 5$ . Algorithm RANDOM GREEDY run on  $H \sim \mathcal{H}(n, k, \Delta)$  produces a matching which covers all but*

$$\left[ \left( \frac{1}{(k-1)(\Delta-1)} \right)^{\frac{\Delta}{(k-1)(\Delta-1)-1}} + o(1) \right] n$$

*many vertices whp.*

---

<sup>1</sup>We say a sequence of events  $A_n$  happens **with high probability** (whp), if  $\lim_{n \rightarrow \infty} \mathbb{P}[A_n] = 1$

In [5], Cooper, Frieze, Molloy and Reed used the small subgraph conditioning method of Robinson and Wormald [11, 12] to prove that

$$\lim_{n \rightarrow \infty} \mathbb{P}[\mathcal{H}(n, k, \Delta) \text{ has a perfect matching}] = \begin{cases} 1 & \text{if } k < \sigma_\Delta \\ 0 & \text{if } k > \sigma_\Delta \end{cases}$$

where  $\sigma_\Delta := \frac{\log \Delta}{(\Delta-1) \log(\frac{\Delta}{\Delta-1})} + 1$ . Thus for any  $k$ ,  $f(k) = \min \{\Delta : k < \sigma_\Delta\}$  gives the threshold of  $\Delta$  such that  $\mathcal{H}(n, k, \Delta)$  has a perfect matching and for large  $k$ ,  $f(k) \sim e^{k-1}$ . Thus it is interesting to note that near this threshold, the greedy algorithm finds a matching which, asymptotically in  $k$ , covers only  $(1 - e^{-1} + o(1))$  fraction of the vertices even though there is a perfect matching whp.

The proof of Theorem 1.1 is an application of the so called differential equations method. In fact we will directly apply a general theorem of Wormald [14] to show concentration of the numbers of vertices of degree  $i$  around their expected trajectories. In Section 2 we will set up the system of  $\Delta$  equations for general fixed degree sequence hypergraphs. In Section 3 we will show how to solve the system and prove Corollary 1.2 using Theorem 1.1. In the last section, we will complete the proof of Theorem 1.1 by applying Wormald's general result.

## 2 Setting up the analysis

Our set up is based heavily on that which appears in [1] where the Karp-Sipser algorithm is analyzed on random graphs with fixed degree sequence. We use the natural extension to hypergraphs of the standard configuration model (see e.g. [15, 9, 2]) to generate  $\mathcal{H}(n, k, \mathbf{z})$ . Each vertex is associated with a set of **configuration points** corresponding to the degree of the vertex. So initially, we will have  $\zeta_i n$  many vertices with  $i$  configuration points and thus  $n \sum_{i=1}^{\Delta} i \zeta_i$  many configuration points total. A **configuration** is a uniform random perfect matching on the collection of configuration points. To obtain a (multi)hypergraph on the original vertex set, we contract all the configuration points corresponding to a vertex into one vertex. The resulting (multi)hypergraph is denoted  $\mathcal{H}^*(n, k, \mathbf{z})$ . A straightforward generalization of the arguments in, say [15], show that this process produces a simple hypergraph with probability bounded away from 0 and conditioning on simplicity, the hypergraph produced is distributed as  $\mathcal{H}(n, k, \mathbf{z})$ . So any result which holds whp in  $\mathcal{H}^*(n, k, \mathbf{z})$  must hold whp in  $\mathcal{H}(n, k, \mathbf{z})$ .

We will only reveal information about the edges in the configuration as the algorithm needs it. The algorithm can be viewed as a sequence of two types of moves on the configuration model. When selecting an edge to add to the matching, the algorithm chooses the  $k$  configuration points uniformly at random which make up the edge. Each such random choice is called a **selection move**. If a selection move is a configuration point of a vertex,  $v$ , of degree  $\ell$ , then the algorithm will delete  $\ell - 1$  other edges. Deleting such an edge will require the random choice of  $k - 1$  configuration points (only  $k - 1$  choices, not  $k$ , since one of them will correspond to the vertex  $v$ ). Each such choice of configuration point will be called a **deletion move**. When we say that “a selection (or deletion) move is a vertex of degree  $i$ ,” we mean that it is a configuration point corresponding to a vertex of degree  $i$ . We will parameterize the algorithm by the number edges in our matching so far (which of course is  $k$  times the number of selection moves). Thus we refer to a **step** as  $k$  selection moves followed by the resulting deletion moves.

The vector  $\mathbf{Y}(j) \in \mathbb{Z}_+^\Delta$  represents the remaining configuration model after  $j$  steps. So there are  $Y_i(j)$  many vertices of degree  $i$  for  $i = 1, \dots, \Delta$ . The vector  $\mathbf{z}(t)$  with  $t = j/n$  is our continuous approximation to  $\mathbf{Y}(j)$ . In other words, we will see later that  $Y_i(j) \approx nz_i(t)$ . Let

$$M(j) = \frac{1}{k} \sum_{i=1}^{\Delta} iY_i(j)$$

So that  $kM(j)$  represents the total number of configuration points remaining after  $j$  steps. For  $i = 1, \dots, \Delta$ , we have

$$\begin{aligned} \mathbb{E}[Y_i(j+1) - Y_i(j) | \mathbf{Y}(j)] &= -k \cdot \frac{iY_i}{kM} \\ &+ k \sum_{\ell=2}^{\Delta} \frac{\ell Y_\ell}{kM} (\ell-1)(k-1) \left( -\frac{iY_i}{kM} + \frac{(i+1)Y_{i+1}}{kM} \right) + O\left(\frac{1}{M}\right). \end{aligned} \quad (2.1)$$

The first term accounts for the event that the selection move is a vertex of degree  $i$ . The next term accounts for the resulting deletion moves. The probability that a selection move is a vertex of degree  $\ell$  is  $\frac{\ell Y_\ell(j) + O(1)}{kM(j) + O(1)} = \frac{\ell Y_\ell}{kM} + O\left(\frac{1}{M}\right)$  (the error terms are due to the fact that during a single step we may lose a few configuration points, bounded by some constant). If the selection is a vertex of degree  $\ell$ , the algorithm makes  $(k-1)(\ell-1)$  deletion moves. If the deletion move is a vertex of degree  $i$  (this happens with probability  $\frac{iY_i + O(1)}{kM + O(1)} = \frac{iY_i}{kM} + O\left(\frac{1}{M}\right)$ ), then we lose a vertex of degree  $i$ , but if the deletion move is a vertex of degree  $i+1$ , then we gain one.

### 3 Analyzing the system of differential equations

For general fixed degree sequences, we are able to reduce the  $\Delta$ -by- $\Delta$  system on  $z_1, \dots, z_\Delta$  to a two-by-two system. However in the general case we cannot explicitly solve that two-by-two system. In the regular case though, we arrive at an explicit solution.

#### 3.1 Solving the system

First we define

$$\begin{aligned} m &:= \frac{1}{k} \sum_{\ell=1}^{\Delta} \ell z_\ell \\ p_\ell &:= \frac{\ell z_\ell}{km}, \quad \ell = 1, \dots, \Delta \\ s &:= (k-1) \sum_{\ell=2}^{\Delta} (\ell-1) p_\ell \end{aligned} \quad (3.1)$$

and

$$z_{\Delta+1} := 0, \quad p_{\Delta+1} := 0.$$

Now if we refer back to (2.1), we see that the system of differential equations modeling our trajectories<sup>2</sup> is

$$\begin{aligned} z'_i &= -k \cdot p_i + k \cdot s \cdot (-p_i + p_{i+1}) \\ &= -k(1+s)p_i + ksp_{i+1}, \quad i = 1, \dots, \Delta \end{aligned} \quad (3.2)$$

with initial conditions

$$z_i(0) = \zeta_i, \quad i = 1, \dots, \Delta. \quad (3.3)$$

Note that this is really a system on only the variables  $z_i$ , since  $s$ ,  $p_i$ , and  $m$  are functions of the  $z_i$ . Substituting the formula for  $p_i$ , we get

$$z'_i + \frac{(1+s)iz_i}{m} = \frac{(i+1)sz_{i+1}}{m} \quad (3.4)$$

In order to follow our solution to this system, it helps to view the above equation as a first order linear differential equation on  $z_i$  (ignoring for the moment the fact that  $s$  and  $m$  also depend on  $z_i$ ). If we define

$$A(t) := \exp \left\{ - \int_0^t \frac{1+s(\tau)}{m(\tau)} d\tau \right\} \quad (3.5)$$

then the integrating factor for the equation (3.4) is  $A^{-i}$ . Multiplying both sides by this integrating factor we arrive at

$$A(t)^{-i} z'_i + A(t)^{-i} \frac{(1+s)iz_i}{m} = [A(t)^{-i} \cdot z_i]' = \frac{A(t)^{-i}(i+1)sz_{i+1}}{m}.$$

Integrating both sides, we get

$$A(t)^{-i} \cdot z_i(t) - \zeta_i = \int_0^t \frac{A(\tau)^{-i}(i+1)s(\tau)z_{i+1}(\tau)}{m(\tau)} d\tau.$$

So we get the recursion

$$z_i(t) = A(t)^i \left[ \zeta_i + \int_0^t \frac{A(\tau)^{-i}(i+1)s(\tau)z_{i+1}(\tau)}{m(\tau)} d\tau \right]$$

which for  $i = \Delta$  immediately gives

$$z_\Delta = A(t)^\Delta \cdot \zeta_\Delta. \quad (3.6)$$

Now if we define

$$B(t) := \int_0^t \frac{s(\tau)A(\tau)}{m(\tau)} d\tau \quad (3.7)$$

then we can write the  $z_i$  in terms of  $A$  and  $B$ .

---

<sup>2</sup>Readers unfamiliar with the differential equations method should refer to the surveys [8] and [14] or more specifically Theorem 4.1 in Section 4 of this paper.

**Claim 3.1.** *We have that*

$$z_i(t) = A(t)^i \sum_{\ell=i}^{\Delta} \zeta_{\ell} \binom{\ell}{i} B(t)^{\ell-i} \quad i = 1, \dots, \Delta. \quad (3.8)$$

*Proof.* We have already shown that the above formula holds for  $i = \Delta$  in (3.6). Now we prove the rest by reverse induction. Assuming (3.8) holds for  $z_{i+1}$ , we get

$$\begin{aligned} z_i &= A^i \left( \zeta_i + \int_0^t \frac{(i+1)z_{i+1}s}{mA^i} d\tau \right) \\ &= A^i \left( \zeta_i + \int_0^t \frac{(i+1)A^{i+1} \sum_{\ell=i+1}^{\Delta} \zeta_{\ell} \binom{\ell}{i+1} B^{\ell-i-1}s}{mA^i} d\tau \right) \\ &= A^i \left( \zeta_i + \sum_{\ell=i+1}^{\Delta} \zeta_{\ell}(i+1) \binom{\ell}{i+1} \int_0^t \frac{B^{\ell-i-1}sA}{m} d\tau \right). \end{aligned} \quad (3.9)$$

Now, since  $B'(t) = \frac{sA}{m}$  and  $B(0) = 0$  (by (3.7)), we see that

$$\int_0^t \frac{B^{\ell-i-1}sA}{m} d\tau = \int_0^t B^{\ell-i-1} B' d\tau = \frac{B^{\ell-i}}{\ell-i}.$$

Thus continuing with (3.9) and using the fact that  $\frac{i+1}{\ell-i} \binom{\ell}{i+1} = \binom{\ell}{i}$ , we have that

$$\begin{aligned} z_i &= A^i \left( \zeta_i + \sum_{\ell=i+1}^{\Delta} \zeta_{\ell}(i+1) \binom{\ell}{i+1} \frac{B^{\ell-i}}{\ell-i} \right) \\ &= A^i \left( \zeta_i + \sum_{\ell=i+1}^{\Delta} \zeta_{\ell} \binom{\ell}{i} B^{\ell-i} \right) \\ &= A(t)^i \sum_{\ell=i}^{\Delta} \zeta_{\ell} \binom{\ell}{i} B(t)^{\ell-i}. \end{aligned}$$

which completes the proof of (3.8).  $\square$

Now that we have written the  $z_i$  in terms of  $A$  and  $B$ , we will reduce the original system (3.4) to a two-by-two system. To do this, we first rewrite  $m$  and  $s$ .

**Claim 3.2.** *We have that*

$$m = \frac{1}{k} AP'(C) \quad (3.10)$$

$$s = \frac{(k-1)AP''(C)}{P'(C)} \quad (3.11)$$

where

$$C = C(t) := A(t) + B(t)$$

and

$$P(x) := \sum_{i=1}^{\Delta} \zeta_i x^i.$$

*Proof.* To prove the formula for  $m$ , we see

$$km = \sum_{\ell=1}^{\Delta} \ell z_{\ell} = \sum_{\ell=1}^{\Delta} \ell A^{\ell} \sum_{i=\ell}^{\Delta} \zeta_i \binom{i}{\ell} B^{i-\ell} = \sum_{i=1}^{\Delta} \sum_{\ell=1}^i A^{\ell} \zeta_i \binom{i}{\ell} B^{i-\ell} = \sum_{i=1}^{\Delta} \sum_{\ell=1}^i A^{\ell} \zeta_i i \binom{i-1}{\ell-1} B^{i-\ell}$$

where we have changed the order of summation and used the fact that  $\ell \binom{i}{\ell} = i \binom{i-1}{\ell-1}$ . Thus, using the binomial theorem and substituting  $A + B = C$ , we have

$$\begin{aligned} km &= \sum_{i=1}^{\Delta} i \zeta_i \sum_{\ell=1}^i \binom{i-1}{\ell-1} A^{\ell} B^{i-\ell} = A \sum_{i=1}^{\Delta} i \zeta_i [A + B]^{i-1} \\ &= A \sum_{i=1}^{\Delta} i \zeta_i C^{i-1} = AP'(C) \end{aligned}$$

and we have proved (3.10). For  $s$ , we have

$$\begin{aligned} s &= (k-1) \sum_{\ell=2}^{\Delta} (\ell-1) p_{\ell} = \frac{k-1}{km} \sum_{\ell=2}^{\Delta} \ell(\ell-1) z_{\ell} \\ &= \frac{k-1}{km} \sum_{\ell=2}^{\Delta} \ell(\ell-1) A^{\ell} \sum_{i=\ell}^{\Delta} \zeta_i \binom{i}{\ell} B^{i-\ell} \\ &= \frac{k-1}{km} \sum_{i=2}^{\Delta} \sum_{\ell=2}^i A^{\ell} \zeta_i \ell(\ell-1) \binom{i}{\ell} B^{i-\ell}. \end{aligned}$$

where we have changed the order of summation. Now using the fact that  $\ell(\ell-1) \binom{i}{\ell} = i(i-1) \binom{i-1}{\ell-2}$ , and the binomial theorem, we have that

$$s = \frac{k-1}{km} \sum_{i=2}^{\Delta} i(i-1) \zeta_i \sum_{\ell=2}^i \binom{i-1}{\ell-2} A^{\ell} B^{i-\ell} = \frac{k-1}{km} A^2 \sum_{i=2}^{\Delta} i(i-1) \zeta_i [A + B]^{i-2}.$$

So substituting (3.10) and  $A + B = C$ , we have

$$\begin{aligned} s &= \frac{(k-1) \left( A^2 \sum_{i=2}^{\Delta} i(i-1) \zeta_i C^{i-2} \right)}{A \sum_{i=1}^{\Delta} i \zeta_i C^{i-1}} \\ &= \frac{(k-1) A \left( \sum_{i=2}^{\Delta} i(i-1) \zeta_i C^{i-2} \right)}{\sum_{i=1}^{\Delta} i \zeta_i C^{i-1}} = \frac{(k-1) AP''(C)}{P'(C)} \end{aligned}$$

which proves (3.11). □

Now recall that by their definitions, (3.5) and (3.7), we have

$$A' = -\frac{(1+s)A}{m}, \quad B' = \frac{sA}{m} = \frac{s(C-B)}{m}, \quad C' = A' + B' = -\frac{A}{m}$$

so substituting (3.10) and (3.11) and simplifying, we have

$$B' = \frac{k(k-1)(C-B)P''(C)}{P'(C)^2} \tag{3.12}$$

$$C' = -\frac{k}{P'(C)}. \quad (3.13)$$

Note that (3.13) only depends on  $C$  and so we can actually solve for  $C$ :

$$P'(C)C' = -k.$$

Integrating both sides and using  $C(0) = 1$  gives

$$P(C) - P(1) = -kt$$

and since  $P(1) = \sum_{i=1}^{\Delta} \zeta_i = 1$  we have

$$P(C) = 1 - kt. \quad (3.14)$$

Now we will solve for  $B$  in terms of  $C$ . Note that (3.12) is equivalent to the first order linear differential equation

$$B' + \frac{k(k-1)P''(C)}{P'(C)^2} \cdot B = \frac{k(k-1)CP''(C)}{P'(C)^2}$$

which, using (3.13), is equivalent to

$$B' - \frac{(k-1)P''(C)C'}{P'(C)} \cdot B = -\frac{(k-1)CP''(C)C'}{P'(C)}$$

which, after multiplying by the integrating factor  $P'(C)^{-(k-1)}$ , becomes

$$\left[ B \cdot P'(C)^{-(k-1)} \right]' = -(k-1)CP'(C)^{-k}P''(C)C'.$$

now integrating both sides and using  $B(0) = 0$ , we get

$$B \cdot P'(C)^{-(k-1)} = Q(C)$$

where

$$Q(x) := \int_x^1 \frac{(k-1)wP''(w)}{P'(w)^k} dw.$$

Thus

$$B = Q(C)P'(C)^{k-1} \quad (3.15)$$

### 3.2 Determining the stopping point

The algorithm stops when there are no more edges. Heuristically, since we expect the number of edges  $M$  to be approximately  $nm$ , we set  $m = 0$ . So using (3.10) the final value of  $t$  should be a root of the equation  $AP'(C) = 0$ . Consider the possibility that  $P'(C) = 0$ .  $C$  is always nonnegative, and  $P'(C)$  is a polynomial with nonnegative coefficients, so we could only have  $P'(C) = 0$  if  $C = 0$ , in which case we would also have  $P(C) = 0$  which by line (3.14) would happen only when  $t = \frac{1}{k}$ . We do not expect this to be the final value of  $t$ , since it would correspond to a perfect matching. Thus we consider the other possibility for the final value of  $t$ : we set  $A = 0$ . Since  $A = C - B$  we would have  $B = C$ . Thus, by (3.15)  $t$  would be a root of the equation

$$C = Q(C)P'(C)^{k-1}$$



(recall that  $C$  is a function of  $t$  given implicitly as the root of  $P(C) = 1 - kt$ ). Let

$$h(x) := \frac{x}{P'(x)^{k-1}} - Q(x).$$

If  $C = Q(C)P'(C)^{k-1}$  then we either have  $C = P'(C) = 0$  (which again can only happen when  $t = \frac{1}{k}$ ) or we have  $h(C) = 0$ . We will show that  $h$  has a unique zero in the interval  $(0, 1)$ , which will then be the value of  $C$  we are really interested in. To see this, first note that by its definition,  $h$  is continuous on  $(0, 1]$  (but possibly not at  $x = 0$  since it is possible that  $P'(0) = 0$ ). Also,

$$h(1) = \frac{1}{P'(1)^{k-1}} > 0. \quad (3.16)$$

Now using the definition of  $Q$ , we have that

$$h'(x) = \frac{1}{P'(x)^{k-1}} - \frac{(k-1)xP''(x)}{P'(x)^k} - Q'(x) = \frac{1}{P'(x)^{k-1}} > 0 \quad (3.17)$$

so  $h$  has at most one zero in the interval  $(0, 1)$ . Now we will be done if we show that

$$\lim_{x \rightarrow 0^+} h(x) < 0$$

(we allow the possibility that the above limit is  $-\infty$ ). This is clear if  $P'(0) \neq 0$ , since then  $h$  is continuous at  $x = 0$  and  $h(0) = -Q(0) < 0$ . So consider the case that  $P'(0) = 0$ . Now from (3.16) and (3.17) we see that

$$h(x) = h(1) - \int_x^1 h'(w) dw = \frac{1}{P'(1)^{k-1}} - \int_x^1 \frac{1}{P'(w)^{k-1}} dw$$

and now it is easy to see that  $\lim_{x \rightarrow 0^+} \int_x^1 \frac{1}{P'(w)^{k-1}} dw = \infty$  since the denominator of the integrand is a polynomial that has a factor  $w$ . Therefore in this case we have  $\lim_{x \rightarrow 0^+} h(x) = -\infty$ . We conclude that  $h$  has a unique zero in the interval  $(0, 1)$ , which we call  $C_{end}$ .

Furthermore, if we let  $t_{end}$  be the value of  $t$  corresponding to  $C_{end}$ , then  $1 - kt_{end} = P(C_{end}) > 0$  so  $t_{end} < \frac{1}{k}$ . Thus,  $t_{end}$  is the smallest value of  $t$  such that  $m = 0$ .

### 3.3 The regular case

In this section we prove Corollary 1.2, given Theorem 1.1.

*Proof of Corollary 1.2.* In this case, we have  $P(x) = x^\Delta$ , and so

$$\begin{aligned} Q(x) &= \int_x^1 \frac{(k-1)wP''(w)}{P'(w)^k} dw \\ &= \int_x^1 \frac{(k-1)(\Delta-1)}{\Delta^{k-1}} w^{-(\Delta-1)(k-1)} dw \\ &= \frac{(k-1)(\Delta-1)}{\Delta^{k-1}[(\Delta-1)(k-1)-1]} \left\{ x^{-[(\Delta-1)(k-1)-1]} - 1 \right\} \end{aligned}$$

where on the last line we have used the fact that  $(\Delta - 1)(k - 1) > 1$  which follows from  $k, \Delta \geq 2$  and  $k + \Delta \geq 5$ . So  $C_{end}$  is the value of  $C$  such that

$$\begin{aligned} C &= Q(C)P'(C)^{k-1} \\ &= \frac{(k-1)(\Delta-1)}{\Delta^{k-1}[(\Delta-1)(k-1)-1]} \left\{ C^{-[(\Delta-1)(k-1)-1]} - 1 \right\} \cdot (\Delta C^{\Delta-1})^{k-1} \\ &= \frac{(k-1)(\Delta-1)}{(\Delta-1)(k-1)-1} \left\{ C - C^{(\Delta-1)(k-1)} \right\} \end{aligned}$$

which can be solved for  $C$  to get

$$C_{end} = \left( \frac{1}{(k-1)(\Delta-1)} \right)^{\frac{1}{(k-1)(\Delta-1)-1}}.$$

Therefore

$$1 - kt_{end} = P(C_{end}) = \left( \frac{1}{(k-1)(\Delta-1)} \right)^{\frac{\Delta}{(k-1)(\Delta-1)-1}}.$$

□

## 4 Showing concentration around the trajectory

We use the following theorem of Wormald [14] as it appears in [8].

**Theorem 4.1.** *Given random variables  $Y_1, \dots, Y_\Delta$  representing components of a time discrete Markov process  $\{G_j\}_{j \geq 0}$ , assume that  $D \subseteq \mathbb{R}^{\Delta+1}$  is closed and bounded and contains the point  $(0, \zeta_1, \dots, \zeta_\Delta)$ . If for all  $j$  and all  $i$ ,  $|Y_i(j+1) - Y_i(j)| \leq \beta$  for some constant  $\beta$  and  $|\mathbb{E}[Y_i(j+1) - Y_i(j)|G_j] - f_i(j/n, Y_1(j/n), \dots, Y_\Delta(j/n))| \leq \lambda$  for some  $\lambda = o(1)$  and some functions  $f_i$  which are Lipschitz continuous on an open set containing  $D$ , then the system of differential equations*

$$\frac{dz_i}{dt} = f_i(t, z_1, \dots, z_\Delta)$$

*has a unique solution  $\hat{z}_1, \dots, \hat{z}_\Delta$  with  $(t, \hat{z}_1(t), \dots, \hat{z}_\Delta(t)) \in D$  satisfying the initial condition  $\hat{z}_i(0) = \zeta_i$ . Moreover, with high probability*

$$Y_i(j) = nz_i(j/n) + o(n)$$

*uniformly for all  $j$  and  $i$ .*

For this application we have

$$f_i = -\frac{iz_i}{km} + \sum_{\ell=2}^{\Delta} \frac{\ell z_\ell}{km} (\ell-1)(k-1) \left( -\frac{iz_i}{km} + \frac{(i+1)z_{i+1}}{km} \right)$$

where

$$m := \frac{1}{k} \sum_{i=1}^{\Delta} iz_i$$

(We continue to use  $m$  to represent the expression above, so for example  $f_i$  as given above is really a function of  $t, z_1, \dots, z_\Delta$ ). We will let the region  $D$  be

$$\left\{ (t, z_1, \dots, z_\Delta) \in \mathbb{R}_+^{\Delta+1} : \quad t \leq \frac{1}{k}, \quad \varepsilon \leq km \leq \sum_{j=1}^{\Delta} i\zeta_i \right\}$$

which is clearly closed and bounded, and contains the point  $(0, \zeta_1, \dots, \zeta_\Delta)$  (where we have  $km = \sum_{j=1}^{\Delta} i\zeta_i$ ). Since  $f_i$  is a rational function whose denominator is only a power of  $m$ ,  $f_i$  is easily seen to be Lipschitz continuous on an open set containing  $D$  (say, the set of points in  $\mathbb{R}_+^{\Delta+1}$  with  $km > \varepsilon/2$ ).

Furthermore, we can see that the only way the solution to the system of differential equations can ever leave  $D$  is at a point where  $km = \sum_{i=1}^{\Delta} iz_i = \varepsilon$ . Indeed, by lines (3.5) and (3.13) we see that  $A$  and  $C$  are decreasing, and the fact that  $A(0) = C(0) = 1$  we have that  $A(t), C(t) \leq 1$  for all  $t$  and so

$$km(t) = A(t) \sum_{i=1}^{\Delta} i\zeta_i C^{i-1}(t) \leq \sum_{i=1}^{\Delta} i\zeta_i.$$

Also, by (3.5) and (3.7) we see that  $A, B \geq 0$ , and then by (3.8) we see that  $z_i \geq 0$  for all  $i$ . Thus, the only inequality defining  $D$  that our solution can ever fail to satisfy is  $\varepsilon \leq km$ .

We are now ready to prove the main theorem.

*Proof of Theorem 1.1.* We apply Theorem 4.1, and conclude that our discrete random variables  $Y_i(j)$  are well approximated by their continuous counterparts  $nz_i(j/n)$ , for all values of  $j \leq t_\varepsilon n$ , where  $t_\varepsilon$  is the value of  $t$  such that  $km(t) = \varepsilon$  (i.e. for values of  $j$  corresponding to points in the region  $D$  defined above). Note that

$$\begin{aligned} km' &= \sum_{i=1}^{\Delta} iz'_i = \sum_{i=1}^{\Delta} i(-k(1+s)p_i + ksp_{i+1}) \\ &= -k \sum_{i=1}^{\Delta} ip_i + ks \sum_{i=1}^{\Delta} (ip_{i+1} - ip_i) \\ &= -k \sum_{i=1}^{\Delta} p_i - ks \sum_{i=1}^{\Delta} p_i \\ &\leq -k \end{aligned}$$

where on the third line we have used the telescoping property of the sum  $\sum_{i=1}^{\Delta} (ip_{i+1} - ip_i)$ . Since  $km(t_{\text{end}}) = 0$  we have

$$t_{\text{end}} - \frac{\varepsilon}{k} \leq t_\varepsilon \leq t_{\text{end}}.$$

We use this to bound the final size of the matching. If we run the process to step  $j_\varepsilon := t_\varepsilon \cdot n$  then by Theorem 4.1 whp we arrive at some configuration with  $kM(j_\varepsilon) = \varepsilon n(1 + o(1))$  many configuration points. At this point our matching already has  $j_\varepsilon = t_\varepsilon n \geq (t_{\text{end}} - \varepsilon)n$  many edges. Also, even if every edge remaining is added to our matching, the final matching will have only  $\varepsilon n(1 + o(1))$  more edges. Thus the maximum possible number of edges is  $(t_{\text{end}} + \varepsilon + o(1))n \leq (t_{\text{end}} + 2\varepsilon)n$ . Altogether the final matching w.h.p. has between  $(t_{\text{end}} - \varepsilon)n$  and  $(t_{\text{end}} + 2\varepsilon)n$  many edges. Since  $\varepsilon > 0$  is arbitrary we are done.  $\square$

## 5 A concluding remark

Recently, Brightwell, Janson and Łuczak [3] analyzed the greedy independent set algorithm on random graphs with specified degree sequences which may depend on  $n$ . They gave a formula for the final size of the independent set produced assuming some conditions on the degree sequence; notably that the first moment  $\sum_{i \geq 0} i \zeta_i$  approaches a finite limit (they sometimes also use the assumption that the second moment  $\sum_{i \geq 0} i^2 \zeta_i$  is bounded). We believe that under analogous assumptions, our results can be extended to random hypergraphs with degree sequences depending on  $n$ .

## References

- [1] T. Bohman and A. Frieze. Karp-Sipser on random graphs with a fixed degree sequence. *Combin. Probab. Comput.*, 20(5):721–741, 2011.
- [2] B. Bollobás. *Random graphs*, volume 73 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, second edition, 2001.
- [3] G. Brightwell, S. Janson, and M. Łuczak. The greedy independent set in a random graph with given degrees. [arXiv:1510.05560](#), 2015.
- [4] P. Chebolu. *Topics in random Graphs*. PhD thesis, Carnegie Mellon University, 2008.
- [5] C. Cooper, A. Frieze, M. Molloy, and B. Reed. Perfect matchings in random  $r$ -regular,  $s$ -uniform hypergraphs. *Combin. Probab. Comput.*, 5(1):1–14, 1996.
- [6] M. Dyer and A. Frieze. Randomized greedy matching. *Random Structures Algorithms*, 2(1):29–45, 1991.
- [7] M. Dyer, A. Frieze, and B. Pittel. The average performance of the greedy matching algorithm. *Ann. Appl. Probab.*, 3(2):526–552, 1993.
- [8] J. Daz and D. Mitsche. The cook-book approach to the differential equation method. *Computer Science Review*, 4(3):129 – 151, 2010.
- [9] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [10] L. Lovász and M. D. Plummer. *Matching theory*. AMS Chelsea Publishing, Providence, RI, 2009. Corrected reprint of the 1986 original [MR0859549].
- [11] R. W. Robinson and N. C. Wormald. Almost all cubic graphs are Hamiltonian. *Random Structures Algorithms*, 3(2):117–125, 1992.
- [12] R. W. Robinson and N. C. Wormald. Almost all regular graphs are Hamiltonian. *Random Structures Algorithms*, 5(2):363–374, 1994.

- [13] G. Tinhofer. A probabilistic analysis of some greedy cardinality matching algorithms. *Annals of Operations Research*, 1(3):239–254.
- [14] N. C. Wormald. Differential equations for random processes and random graphs. *Ann. Appl. Probab.*, 5(4):1217–1235, 1995.
- [15] N. C. Wormald. Models of random regular graphs. In *Surveys in combinatorics, 1999 (Canterbury)*, volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.